8-1-2002

Gradient Recovery and A Posteriori Estimate for Bilinear Element on Irregular Quadrilateral Meshes

Zhimin Zhang
Wayne State University, zhimin.zhang@wayne.edu

GRADIENT RECOVERY AND A POSTERIORI ESTIMATE FOR BILINEAR ELEMENT ON IRREGULAR QUADRILATERAL MESHES

ZHIMIN ZHANG

WAYNE STATE UNIVERSITY

Detroit, MI 48202

Department of Mathematics

Research Report

2002 Series
#8

This research was partly supported by the National Science Foundation.

Gradient Recovery and *A Posteriori* Estimate for Bilinear Element on Irregular Quadrilateral Meshes

Zhimin Zhang*

Abstract. A polynomial preserving gradient recovery method is proposed and analyzed for bilinear element under general quadrilateral meshes. It has been proven that the recovered gradient converges at a rate $O(h^{1+\rho})$ for $\rho = \min(\alpha, 1)$ when the mesh is distorted $O(h^{1+\alpha})$ $(\alpha > 0)$ from a regular one. Consequently, the *a posteriori* error estimator based on the recovered gradient is asymptotically exact.

2000 *Mathematics Subject Classification.* Primary 65N30, Secondary 65N15, 41A10, 41A25, 41A27, 41A63.

Key Words and phrases. Finite element method, quadrilateral mesh, gradient recovery, superconvergence, *a posteriori* error estimate.

1. Introduction.

A posteriori error estimate is an active research area and many methods have been developed. Roughly speaking, there are residual type error estimators and recovery type estimators. For the literature, readers are referred to recent books by Ainsworth-Oden [2] and by Babuška-Strouboulis [4], a conference proceeding [16], a survey article by Bank [5], an earlier book by Verfürth [21], and references therein.

While the residual type estimators have been analyzed extensively, there is only a limited theoretical research on recovery type error estimators (see, e.g., [2, Chapter 4], [9, 6, 7, 14, 15, 26]). Yet, the recovery type error estimators are widely used in engineering applications and their practical effectiveness has been recognized by more and more researchers. Currently, ZZ patch recovery error estimator is used in commercial codes, such as ANSYS, MCS/NASTRAN-Marc, Pro/MECHANICA (a product of Parametric Technology), and I-DEAS (product of SDRC, part of EDS), for the purpose of smoothing and adaptive re-meshing. It is also used in NASA's COMET-AR (COmputational MEchanics Testbed With Adaptive Refinement). In a computer based investigation [4] by Babuška et al., it was found that out of all error estimators tested (including the equilibrated residual error estimator and the ZZ patch recovery error estimator and many others), the ZZ patch recovery error estimator based on the discrete least-squares fitting is the most robust.

It is worth pointing out that the recovery type error estimator was originally based on the finite element superconvergence theory, in hopes that a recovered gradient is supercon-

*This research was partially supported by the National Science Foundation grants DMS-0074301, DMS-0079743, and INT-0196139.

vergent and hence can be used as a substitute of the exact gradient to measure the error. The reader is referred to [4, 10, 16, 18, 23, 30] for the literature regarding the superconvergence theory. In order to prove superconvergence, it is necessary to impose some strong restrictions on mesh and these restrictions are usually not satisfied in practice. Nevertheless, it is found that in many practical situations, recovery type error estimators perform astonishingly well under meshes produced by the Delaunay triangulation. Mathematically, this fact has not yet been rigorously justified.

In a recent work, Bank-Xu [6, 7] introduced a recovery type error estimator based on global L_2-projection with smoothing iteration of the multigrid method and established asymptotic exactness in the H^1-norm for linear element under shape regular triangulation. However, the recovery operator is a global one.

For element-wise recovery operator, Hoffmann et al. [14, 15] proposed a general framework which request that, for a fixed $0 < \epsilon < 1$,

$$m = C\left(\left(\frac{H}{h}\right)^2 h^\epsilon + \left(\frac{H}{h}\right)^{-\epsilon} \ln \frac{H}{h}\right) < 1.$$

Here h is the size of element τ, $H \geq 2h$ is the size of the patch ω_τ (surrounding the element), where the recovery takes place, and C is an unknown constant which comes from the analysis. Let $H = Lh$, in order for $m < 1$, we need

$$C(L^2 h^\epsilon + L^{-\epsilon} \ln L).$$

Depending on C, this essentially ask for sufficiently large L and sufficiently small h, which implies many elements may be needed for the recovery procedure.

Therefore a theoretical justification for the recovery that involves only a few elements surrounding a node is necessary. The situation is further complicated by quadrilateral meshes where mappings between the reference element and physical elements are not affine. We encounter some delicate theoretical issue in analysis. See [1, 8, 12, 3, 13, 19, 20, 27, 28, 33] for more details.

In this article, we propose and analyze a gradient recovery method which is different from the ZZ recovery [31]. We show that the *a posteriori* estimate based on this new recovery operator is asymptoticly exact under the mesh distortion $O(h^\alpha)$ when $\alpha > 0$. Here $\alpha = \infty$ represents the uniform mesh and $\alpha = 0$ represents completely unstructured mesh.

The main feature of this new recovery operator is: (1) It is completely local just as the ZZ patch recovery; (2) It is polynomial preserving without any mesh assumption, a property not shared by the ZZ; (3) It is superconvergent under "mildly" constructed grid; (4) It results in an asymptotically exact error estimator when the mesh is not too much distorted. The error bound is in form of

$$\eta_h + O(h^{1+\rho}) \leq \|\nabla(u - u_h)\| \leq \eta_h + O(h^{1+\rho}), \tag{1.1}$$

comparing with

$$\frac{1}{C}\eta_h + \text{higher order term} \leq \|\nabla(u - u_h)\| \leq C\eta_h + \text{higher order term}$$

in most of error bounds in the literature, where C is an unknown constant, which may be very large and hence makes the error bound not very meaningful in practice.

We comment that h^α can be reduced to $o(1)$ and still maintaining the asymptotic exactness of the error estimator. If we give up the asymptotic exact requirement and only ask for a reasonable error estimator, we may further reduce the condition to "a sufficiently small constant $\gamma > 0$".

In order to establish (1.1), a fundamental analysis needs to be performed. The analysis involves some geometric properties of quadrilaterals. For the counterpart of this work under linear triangular element, see [25] by Xu-Zhang.

2. Geometry of the Quadrilateral.

Let $\hat{K} = [-1,1] \times [-1,1]$ be the reference element with vertices $\hat{Z}_i$, and let K be a convex quadrilateral with vertices $Z_i^K(x_i^K, y_i^K)$, $i = 1,2,3,4$. There exists a unique bilinear mapping F_K such that $F_K(\hat{K}) = K, F_K(\hat{Z}_i) = Z_i^K$ given by

$$x = \sum_{i=1}^{4} x_i^K N_i, \quad y = \sum_{i=1}^{4} y_i^K N_i,$$

where

$$N_1 = \frac{1}{4}(1-\xi)(1-\eta) + \frac{1}{4}(1+\xi)(1-\eta) + \frac{1}{4}(1+\xi)(1+\eta) + \frac{1}{4}(1-\xi)(1+\eta).$$

We can also express

$$x = a_0 + a_1\xi + a_2\eta + a_3\xi\eta, \quad y = b_0 + b_1\xi + b_2\eta + b_3\xi\eta;$$

where by suppressing the index "K",

$$\begin{aligned} 4a_0 &= x_1 + x_2 + x_3 + x_4, & 4b_0 &= y_1 + y_2 + y_3 + y_4; \\ 4a_1 &= -x_1 + x_2 + x_3 - x_4, & 4b_1 &= -y_1 + y_2 + y_3 - y_4; \\ 4a_2 &= -x_1 - x_2 + x_3 + x_4, & 4b_2 &= -y_1 - y_2 + y_3 + y_4; \\ 4a_3 &= x_1 - x_2 + x_3 - x_4, & 4b_3 &= y_1 - y_2 + y_3 - y_4. \end{aligned}$$

To any function $v(x,y)$ defined on K, we associate $\hat{v}(\xi,\eta)$ by

$$\hat{v}(\xi,\eta) = v(x(\xi,\eta), y(\xi,\eta)), \quad \text{or} \quad \hat{v} = v \circ F_K.$$

The Jacobi matrix of the mapping F_K is

$$(DF_K)(\xi,\eta) = \begin{pmatrix} x_\xi & y_\xi \\ x_\eta & y_\eta \end{pmatrix} = \begin{pmatrix} a_1 + a_3\eta & b_1 + b_3\eta \\ a_2 + a_3\xi & b_2 + b_3\xi \end{pmatrix}.$$

Let $\nabla v = (\partial_x v, \partial_y v)^T$, it is straight forward to verify that

$$\hat{\nabla}\hat{v} = (\partial_\xi \hat{v}, \partial_\eta \hat{v})^T = DF_K \nabla v, \tag{2.1}$$

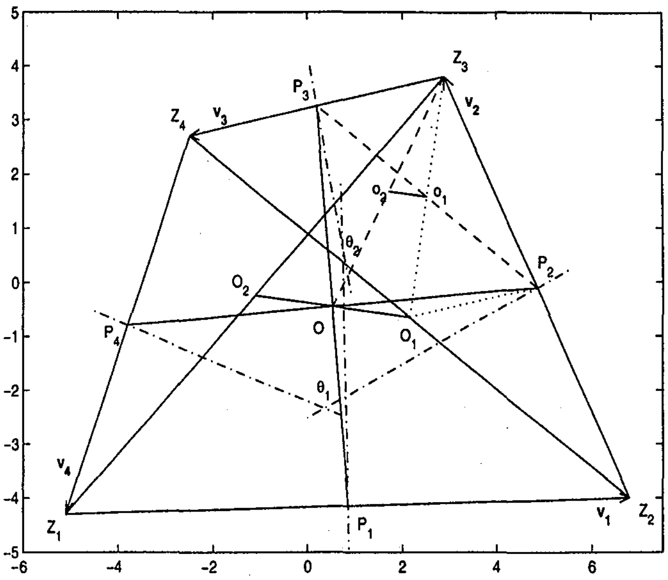

Figure 1: Geometry of a quadrilateral

$$\partial^r_{\xi^r}\hat{v} = [(a_1 + a_3\eta)\partial_x + (b_1 + b_3\eta)\partial_y]^r v, \tag{2.2}$$

$$\begin{aligned}\partial^{r+1}_{\xi^r\eta}\hat{v} &= r[(a_1 + a_3\eta)\partial_x + (b_1 + b_3\eta)\partial_y]^{r-1}(a_3, b_3)\cdot\nabla v \\ &\quad +[(a_1 + a_3\eta)\partial_x + (b_1 + b_3\eta)\partial_y]^r[(a_2 + a_3\xi)\partial_x + (b_2 + b_3\xi)\partial_y]v,\end{aligned} \tag{2.3}$$

and $\partial^r_{\eta^r}\hat{v}$ and $\partial^{r+1}_{\xi\eta^r}\hat{v}$ can be expressed in a similar way. The determinant of the Jacobi matrix is

$$J_K = J_K(\xi, \eta) = J_0^K + J_1^K\xi + J_2^K\eta,$$

where

$$J_0^K = a_1b_2 - b_1a_2, \quad J_1^K = a_1b_3 - b_1a_3, \quad J_2^K = b_2a_3 - a_2b_3.$$

The inverse of the Jacobi matrix is

$$\begin{pmatrix} \xi_x & \eta_x \\ \xi_y & \eta_y \end{pmatrix} = (DF_K)^{-1} = \frac{1}{J_K}\begin{pmatrix} b_2 + b_3\xi & -b_1 - b_3\eta \\ -a_2 - a_3\xi & a_1 + a_3\eta \end{pmatrix}.$$

Note that $a_3 = b_3 = 0$ when K is a parallelogram in which case F_K is an affine mapping, and further $a_3 = b_3 = a_2 = b_1 = 0$ when K is a rectangle.

Starting from Z_1, we express the four edges (with the midpoint P_i) as four vectors $\boldsymbol{v}_i$, $i = 1, 2, 3, 4$, pointing counter-clock-wisely (Figure 1). We denote the midpoints of Z_2Z_4

and Z_1Z_3 as O_1 and O_2, respectively. We can verify that

$$\begin{aligned}
P_4P_2 = \frac{1}{2}(x_2+x_3-x_4-x_1, y_2+y_3-y_4-y_1, 0) &= 2(a_1, b_1, 0), \\
P_1P_3 = \frac{1}{2}(x_3+x_4-x_1-x_2, y_3+y_4-y_1-y_2, 0) &= 2(a_2, b_2, 0), \\
O_1O_2 = \frac{1}{2}(x_1+x_3-x_2-x_4, y_1+y_3-y_2-y_4, 0) &= 2(a_3, b_3, 0).
\end{aligned}$$

Then

$$2\sqrt{a_1^2+b_1^2} = |P_4P_2|, \quad 2\sqrt{a_2^2+b_2^2} = |P_1P_3|, \quad 2\sqrt{a_3^2+b_3^2} = |O_1O_2|. \tag{2.4}$$

$$4(a_1a_2+b_1b_2) = P_4P_2 \cdot P_1P_3 = |P_4P_2||P_1P_3|\cos\alpha_K, \tag{2.5}$$
$$4(a_1a_3+b_1b_3) = P_4P_2 \cdot O_1O_2 = |P_4P_2||O_1O_2|\cos\beta_K, \tag{2.6}$$
$$4(a_2a_3+b_2b_3) = O_1O_2 \cdot P_1P_3 = |O_1O_2||P_1P_3|\cos\gamma_K, \tag{2.7}$$

where the meaning of angles α_K, β_K, and γ_K is obvious from the context.

$$J_0^K\boldsymbol{k} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ a_1 & b_1 & 0 \\ a_2 & b_2 & 0 \end{vmatrix} = \frac{1}{4}P_4P_2 \times P_1P_3 = \frac{1}{4}|P_4P_2||P_1P_3|\sin\alpha_K, \tag{2.8}$$

$$J_1^K\boldsymbol{k} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ a_1 & b_1 & 0 \\ a_3 & b_3 & 0 \end{vmatrix} = \frac{1}{4}P_4P_2 \times O_1O_2 = \frac{1}{4}|P_4P_2||O_1O_2|\sin\beta_K, \tag{2.9}$$

$$J_2^K\boldsymbol{k} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ a_2 & b_2 & 0 \\ a_3 & b_3 & 0 \end{vmatrix} = \frac{1}{4}P_1P_3 \times O_1O_2 = \frac{1}{4}|P_1P_3||O_1O_2|\sin\gamma_K. \tag{2.10}$$

We could also express

$$|J_1^K| = 2|(x_4-x_3)(y_2-y_1)-(x_2-x_1)(y_4-y_3)| = 2|\boldsymbol{v}_3 \times \boldsymbol{v}_1|, \quad |J_2^K| = 2|\boldsymbol{v}_4 \times \boldsymbol{v}_2|$$

Let h_K be the longest edge length of K, we introduce the following condition:
Definition 1. A convex quadrilateral K is said to satisfy the *diagonal condition* if

$$d_K = |O_1O_2| = O(h_K^{1+\alpha}), \quad \alpha \geq 0. \tag{2.11}$$

Note that K is a parallelogram if and only if $d_K = 0$. Therefore, the distance between the two diagonal mid-points O_1 and O_2 is a convenient measure for the deviation of a quadrilateral from a parallelogram. The two extremal cases $\alpha \to \infty$ and $\alpha \to 0$ represent parallelogram and completely unstructured quadrilateral, respectively. Anything in between will pose some restriction, especially $\alpha = 1$ is the well-known 2-strongly regular partition, see, e.g., [12, 33].

The diagonal condition was previously used by Chen [11] for triangular meshes, where two adjacent triangles form a quadrilateral that satisfies the condition.

The following lemma states a known fact regarding the 2-strongly regular partition ($\alpha = 1$). Although this fact is widely used, we have not seen a formal proof of it in the literature. An elementary proof is therefore provided in the Appendix.

Lemma 2.1. Let o_1o_2 be the distance between two diagonal mid-points of any of four refined quadrilaterals through the bi-section of K. Then

$$|o_1o_2| = \frac{1}{4}|O_1O_2|.$$

Recall that the bi-section reduces the length of longest edge by half, which is $h_K/2$. Therefore, the diagonal condition (2.11) is satisfied with $\alpha = 1$.

To measure this deviation, Rannarchar and Turek [20] used the quantity

$$\sigma_K = \max(|\pi - \theta_1|, |\pi - \theta_2|),$$

where θ_1 and θ_2 are the angles between the outward normals of two opposite sides of K.

Definition 2. A convex quadrilateral K is said to satisfy the *angle condition* if

$$\sigma_K = O(h_K^\alpha), \quad \alpha \geq 0. \tag{2.12}$$

Lemma 2.2. The diagonal condition (2.11) and the angle condition (2.12) are equivalent in the sense

$$d_K = O(h_K^{1+\alpha}) \Longleftrightarrow \sigma_K = O(h_K^\alpha), \quad \alpha \geq 0.$$

A special case of this lemma has been proved in [19, Theorem 4.13] under some complicated mesh restrictions. Here we provide a direct and much simpler proof in the Appendix without any mesh assumption.

Definition 3. A partition $\mathcal{T}_h$ is said to satisfy *Condition* (α) if there exist $\alpha > 0$ such that

i) Any $K \in \mathcal{T}_h$ satisfies the diagonal condition (2.11).

ii) Any two K_1, K_2 in $\mathcal{T}_h$ that share a common edge satisfy a *neighboring condition,*

$$a_j^{K_1} = a_j^{K_2}(1 + O(h_{K_1}^\alpha + h_{K_2}^\alpha)), \quad b_j^{K_1} = b_j^{K_2}(1 + O(h_{K_1}^\alpha + h_{K_2}^\alpha)), \quad j = 1, 2. \tag{2.13}$$

To assure optimal order error estimates in the H^1-norm for the bilinear isoparametric interpolation on a convex quadrilateral K, namely, the estimate

$$\|u - u_I\|_{0,K} + h|u - u_I|_{1,K} \leq Ch_K^2|u|_{2,K}, \tag{2.14}$$

we need a degeneration condition, which was introduced by Acosta and Durán [1].

Definition 4. A convex quadrilateral K is said to satisfy the *Regular decomposition property* with constants $N \in R$ and $0 < \Psi < \pi$, or shortly $RDP(N, \Psi)$, if we can divide K into two triangles along one of its diagonals, which will always be called d_1, in such a way that $|d_1|/|d_2| \leq N$ and both triangles satisfy the *maximum angle condition* with parameter Ψ (i.e., all angles are bounded by Ψ).

Remark. This is a weaker condition than many other similar degenerate conditions, cf. e.g., [12, 13, 28, 33]. It was proved in [1] that $RDP(N,\Psi)$ is a sufficient condition for (2.14) to be hold, and the authors conjectured that it is also a necessary condition. Recently, Ming-Shi confirmed this conjecture by a simple counter-example [19].

We denote $X = X(\xi,\eta) = X_0 + X_1$ where

$$X_0 = \begin{pmatrix} b_2 & -b_1 \\ -a_2 & a_1 \end{pmatrix}, \quad X_1 = X_1(\xi,\eta) = \begin{pmatrix} b_3 \\ -a_3 \end{pmatrix}(\xi,-\eta).$$

Lemma 2.3. Let a convex quadrilateral K satisfy the diagonal condition. Then

$$\|X_0X^{-1}\|_2 = 1 + O(h_K^\alpha), \quad \|X_1X^{-1}\|_2 = \|I - X_0X^{-1}\|_2 = O(h_K^\alpha).$$

Proof: It is straightforward to verify that

$$\begin{aligned} X_0X^{-1} &= \begin{pmatrix} b_2 & -b_1 \\ -a_2 & a_1 \end{pmatrix} \frac{1}{J_K}\left[\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} + \begin{pmatrix} \eta \\ \xi \end{pmatrix}(a_3,b_3)\right] \\ &= \frac{J_0^K}{J_K}I + \frac{1}{J_K}\begin{pmatrix} b_2 & -b_1 \\ -a_2 & a_1 \end{pmatrix}\begin{pmatrix} \eta \\ \xi \end{pmatrix}(a_3,b_3) \end{aligned}$$

where I is a 2-by-2 identity matrix; and

$$X_1X^{-1} = I - X_0X^{-1} = (\frac{J_1^K}{J_K}\xi + \frac{J_2^K}{J_K}\eta)I - \frac{1}{J_K}\begin{pmatrix} b_2 & -b_1 \\ -a_2 & a_1 \end{pmatrix}\begin{pmatrix} \eta \\ \xi \end{pmatrix}(a_3,b_3).$$

By the definition of J_K and geometric relations of (2.4), (2.8)–(2.10), we see that

$$\frac{J_0^K}{J_K} = 1 + O(h_K^\alpha), \quad \frac{J_1^K}{J_K} = O(h_K^\alpha), \quad \frac{J_2^K}{J_K} = O(h_K^\alpha),$$

by the diagonal condition (2.11). The desired conclusion follows. □

3. Superconvergence Analysis.

We consider the variational problem: Find $u \in H^1(\Omega)$ such that

$$a(u,v) = (\nabla u, A\nabla v) + (\boldsymbol{b}\cdot\nabla u, v) + (cu,v) = (f,v), \quad \forall v \in H^1(\Omega), \tag{3.1}$$

where A is a 2-by-2 symmetric positive definite matrix and Ω is a polygonal domain which allows a quadrilateral partition $\mathcal{T}_h$ with $h = \max_{K\in\mathcal{T}_h} h_K$. We assume that all functions are sufficiently smooth, in particular,

$$\|A - A_0\|_{0,\infty,K} = O(h_K^\alpha), \quad \|\boldsymbol{b} - \boldsymbol{b}_0\|_{0,\infty,K} = O(h_K^\alpha), \tag{3.2}$$

where A_0 and $\boldsymbol{b}_0$ are piece-wisely constant functions that on each $K \in \mathcal{T}_h$,

$$A_0|_K = \frac{1}{|K|}\int_K A(x,y)dxdy, \quad \boldsymbol{b}_0|_K = \frac{1}{|K|}\int_K \boldsymbol{b}(x,y)dxdy.$$

We also assume that $a(\cdot,\cdot)$ satisfies the inf-sup condition to insure that (3.1) has a unique solution. Using

$$\nabla v = \frac{1}{J_K} X\hat{\nabla}\hat{v},$$

we write

$$(\nabla w, A\nabla v)_K = \int_K (\nabla w)^T A \nabla v dxdy = \int_{\hat{K}} \frac{1}{J_K} (X\hat{\nabla}\hat{w})^T \hat{A}(X\hat{\nabla}\hat{v}) d\xi d\eta,$$

$$(\boldsymbol{b}\cdot\nabla w, v)_K = \int_K v\boldsymbol{b}\cdot\nabla w dxdy = \int_{\hat{K}} \hat{v}\hat{\boldsymbol{b}}\cdot X\hat{\nabla}\hat{w} d\xi d\eta;$$

and define

$$(\nabla w, A\nabla v)_K^* = \frac{1}{J_0^K}\int_{\hat{K}} (X_0\hat{\nabla}\hat{w})^T A_0 (X_0\hat{\nabla}\hat{v}) d\xi d\eta = \int_{\hat{K}} (\hat{\nabla}\hat{w})^T B^K \hat{\nabla}\hat{v} d\xi d\eta, \tag{3.3}$$

$$(\boldsymbol{b}\cdot\nabla w, v)_K^* = \boldsymbol{b}_0\cdot X_0 \int_{\hat{K}} \hat{v}\hat{\nabla}\hat{w} d\xi d\eta, \tag{3.4}$$

where

$$B^K = \frac{1}{J_0^K}(X_0^K)^T A_0^K X_0^K.$$

We introduce the following lemma, which can be verified by straightforward calculation.

Lemma 3.1. Under the condition (2.13) and (3.2), we have

$$J_0^{K_1} = J_0^{K_2}(1 + O(h_{K_1}^\alpha + h_{K_2}^\alpha)), \quad \|B^{K_1} - B^{K_2}\| = O(h_{K_1}^\alpha + h_{K_2}^\alpha).$$

Theorem 3.1. Let K satisfy the diagonal condition, and the assumption (3.2) be satisfied. Then there exists a constant C independent of u and K, such that

$$|(\nabla w, A\nabla v)_K - (\nabla w, A\nabla v)_K^*| \le C h_K^\alpha \|\nabla w\|_{0,K}\|\nabla v\|_{0,K}, \tag{3.5}$$

$$|(\boldsymbol{b}\cdot\nabla w, v)_K - (\boldsymbol{b}\cdot\nabla w, v)_K^*| \le C h_K^\alpha \|\nabla w\|_{0,K}\|v\|_{0,K}. \tag{3.6}$$

Proof: We decompose

$$\begin{aligned}
& (\nabla w, A\nabla v)_K - (\nabla w, A\nabla v)_K^* \\
= \; & (\nabla w, (A - A_0)\nabla v)_K + \int_{\hat{K}} \frac{1}{J_K}[(X\hat{\nabla}\hat{w})^T A_0 (X\hat{\nabla}\hat{v}) - (X_0\hat{\nabla}\hat{w})^T A_0 (X_0\hat{\nabla}\hat{v})] d\xi d\eta \\
& + \int_{\hat{K}} (\frac{1}{J_K} - \frac{1}{J_0^K})(X_0\hat{\nabla}\hat{w})^T A_0 (X_0\hat{\nabla}\hat{v}) d\xi d\eta.
\end{aligned} \tag{3.7}$$

By (3.2)

$$|(\nabla w, (A - A_0)\nabla v)_K| \le C h_K^\alpha \|\nabla w\|_{0,K}\|\nabla v\|_{0,K}. \tag{3.8}$$

Using $X = X_0 + X_1$, we express

$$\begin{aligned}
&\int_{\hat{K}} \frac{1}{J_K}[(X\hat{\nabla}\hat{w})^T A_0 (X\hat{\nabla}\hat{v}) - (X_0\hat{\nabla}\hat{w})^T A_0 (X_0\hat{\nabla}\hat{v})] d\xi d\eta \\
= &\int_{\hat{K}} \frac{1}{J_K}[(X_0\hat{\nabla}\hat{w})^T A_0 (X_1\hat{\nabla}\hat{v}) + (X_1\hat{\nabla}\hat{w})^T A_0 (X_0\hat{\nabla}\hat{v}) + (X_1\hat{\nabla}\hat{w})^T A_0 (X_1\hat{\nabla}\hat{v})] d\xi d\eta.
\end{aligned}$$

The first term can be estimated as

$$\begin{aligned}
&|\int_{\hat{K}} \frac{1}{J_K}(X_0\hat{\nabla}\hat{w})^T A_0 (X_1\hat{\nabla}\hat{v}) d\xi d\eta| \\
= &|\int_{\hat{K}} (\frac{1}{J_K} X\hat{\nabla}\hat{w})^T X^{-T} X_0^T A_0 X_1 X^{-1} (\frac{1}{J_K} X\hat{\nabla}\hat{v}) J_K d\xi d\eta| \\
= &|\int_K (\nabla w)^T (X_0 X^{-1})^T A_0 X_1 X^{-1} \nabla v dx dy| \leq C h_K^\alpha \|\nabla w\|_{0,K} \|\nabla v\|_{0,K}.
\end{aligned}$$

Note that $(X_0 X^{-1})^T A_0 X_1 X^{-1} = O(h_K^\alpha)$ by Lemma 2.3. The other two terms can be estimated similarly. Then we derive

$$|\int_{\hat{K}} \frac{1}{J_K}[(X\hat{\nabla}\hat{w})^T A_0 (X\hat{\nabla}\hat{v}) - (X_0\hat{\nabla}\hat{w})^T A_0 (X_0\hat{\nabla}\hat{v})] d\xi d\eta| \leq C h_K^\alpha \|\nabla w\|_{0,K} \|\nabla v\|_{0,K}. \quad (3.9)$$

Next

$$\begin{aligned}
&|\int_{\hat{K}} (\frac{1}{J_K} - \frac{1}{J_0^K})(X_0\hat{\nabla}\hat{w})^T A_0 (X_0\hat{\nabla}\hat{v}) d\xi d\eta| \\
= &|\int_{\hat{K}} (1 - \frac{J_K}{J_0^K})(\frac{1}{J_K} X\hat{\nabla}\hat{w})^T X^{-T} X_0^T A_0 X_0 X^{-1} (\frac{1}{J_K} X\hat{\nabla}\hat{v}) J_K d\xi d\eta| \\
= &|\int_K (-\frac{J_1^K}{J_0^K}\xi(x,y) - \frac{J_2^K}{J_0^K}\eta(x,y))(\nabla w)^T (X_0 X^{-1})^T A_0 X_0 X^{-1} \nabla v dx dy| \\
\leq &C h_K^\alpha \|\nabla w\|_{0,K} \|\nabla v\|_{0,K}.
\end{aligned} \quad (3.10)$$

Note that by Lemma 2.3,

$$\frac{J_1^K}{J_0^K}\xi(x,y) + \frac{J_2^K}{J_0^K}\eta(x,y) = O(h_K^\alpha), \quad (X_0 X^{-1})^T A_0 X_0 X^{-1} = 1 + O(h_K^\alpha).$$

We then obtain (3.5) by applying (3.8)-(3.10) to the right hand side of (3.7).

Now we write the convection term as following:

$$\begin{aligned}
&(\boldsymbol{b}\cdot\nabla w, v)_K - (\boldsymbol{b}\cdot\nabla w, v)_K^* \\
= &\int_{\hat{K}} \hat{v}\hat{\boldsymbol{b}}\cdot(X - X_0)\hat{\nabla}\hat{w} d\xi d\eta + \int_{\hat{K}} \hat{v}(\hat{\boldsymbol{b}} - \boldsymbol{b}_0)\cdot X_0\hat{\nabla}\hat{w} d\xi d\eta.
\end{aligned} \quad (3.11)$$

We estimate the two terms separately.

$$
\begin{aligned}
|\int_{\hat{K}} \hat{v}\hat{\boldsymbol{b}} \cdot (X - X_0)\hat{\nabla}\hat{w}d\xi d\eta| &= |\int_{\hat{K}} \hat{v}\hat{\boldsymbol{b}} \cdot (I - X_0X^{-1})X\hat{\nabla}\hat{w}d\xi d\eta| \\
&= |\int_K v\boldsymbol{b} \cdot (I - X_0X^{-1})\nabla w dxdy| \\
&\leq Ch_K^{\alpha}\|\nabla w\|_{0,K}\|v\|_{0,K}, \qquad (3.12)
\end{aligned}
$$

by Lemma 2.3.

$$
\begin{aligned}
|\int_{\hat{K}} \hat{v}(\hat{\boldsymbol{b}} - \boldsymbol{b}_0) \cdot X_0\hat{\nabla}\hat{w}d\xi d\eta| &= |\int_{\hat{K}} \hat{v}(\hat{\boldsymbol{b}} - \boldsymbol{b}_0) \cdot X_0X^{-1}(X\hat{\nabla}\hat{w})d\xi d\eta| \\
&= |\int_K v(\boldsymbol{b} - \boldsymbol{b}_0) \cdot X_0X^{-1}\nabla w dxdy| \\
&\leq Ch_K^{\alpha}\|\nabla w\|_{0,K}\|v\|_{0,K}, \qquad (3.13)
\end{aligned}
$$

by Lemma 2.3 and (3.2). Applying (3.12) and (3.13) to (3.11), we obtain (3.6). □

We then define two modified bilinear forms

$$
a_h(u,v) = \sum_K a_h(u,v)_K, \quad b_h(w,v) = \sum_K b_h(u,v)_K
$$

where

$$
a_h(u,v)_K = (\nabla u, A\nabla v)_K^* + (\boldsymbol{b} \cdot \nabla u, v)_K + (cu, v)_K, \qquad (3.14)
$$

$$
b_h(u,v)_K = (\nabla u, A\nabla v)_K^* + (\boldsymbol{b} \cdot \nabla w, v)_K^* + (cu, v)_K. \qquad (3.15)
$$

Given a quadrilateral partition $\mathcal{T}_h$ on a polygonal domain Ω, we define the bilinear finite element space

$$
S_h = \{v \in H^1(\Omega) : \hat{v} = v \circ F_K \in Q_1(\hat{K}), \ \forall K \in \mathcal{T}_h\}.
$$

Theorem 3.2. Let $\mathcal{T}_h$ satisfy the condition (α) and $RDP(N, \Psi)$, and let $u_I \in S_h$ be the bilinear interpolation of $u \in H^3(\Omega) \cap H_0^1(\Omega)$. Then there exists a constant C independent of h and u, such that for any $v \in S_h$,

$$
|a_h(u - u_I, v)| + |b_h(u - u_I, v)| \leq C(h^{1+\alpha}|u|_{2,\Omega} + h^2|u|_{3,\Omega})\|v\|_{1,\Omega}.
$$

Proof: For convenience, we set $w = u - u_I$. By (3.14) and (3.15), we can express

$$
a_h(w,v) - b_h(w,v) = \sum_{K \in \mathcal{T}_h} [(\boldsymbol{b} \cdot \nabla w, v)_K - (\boldsymbol{b} \cdot \nabla w, v)_K^*].
$$

Recall (3.6), and we have

$$
|a_h(w,v) - b_h(w,v)| \leq C \sum_{K \in \mathcal{T}_h} h_K^{\alpha}\|\nabla w\|_{0,K}\|v\|_{0,K} \leq Ch^{1+\alpha}|u|_{2,\Omega}\|v\|_{0,\Omega}, \qquad (3.16)
$$

since by the $RDP(N,\Psi)$ assumption, (2.14) is valid. Therefore, we only need to estimate $b_h(w,v)$. Again, by the $RDP(N,\Psi)$ assumption, we have

$$|(cw,v)| \leq Ch^2|u|_{2,\Omega}\|v\|_{0,\Omega}. \tag{3.17}$$

Hence, our task is narrowed down to estimate

$$(\nabla u, A\nabla v)_K^*, \quad \text{and} \quad (\boldsymbol{b}\cdot\nabla w, v)_K^*$$

for $K \in \mathcal{T}_h$. By the definition (3.3) and (3.4), we see that all coefficients are constants now and we only need to estimate following terms

$$\int_{\hat{K}} \partial_\xi \hat{w}\partial_\xi \hat{v}, \quad \int_{\hat{K}} \partial_\xi \hat{w}\partial_\eta \hat{v}, \quad \int_{\hat{K}} \partial_\eta \hat{w}\partial_\xi \hat{v}, \quad \int_{\hat{K}} \partial_\eta \hat{w}\partial_\eta \hat{v}, \quad \int_{\hat{K}} \hat{v}\partial_\xi \hat{w}, \quad \int_{\hat{K}} \hat{v}\partial_\eta \hat{w}.$$

a) Let $\hat{u} \in P_2(\hat{K})$. There are only two terms ξ^2, η^2 not in the reference space of the bilinear interpolation, therefore,

$$\int_{\hat{K}} \partial_\xi \hat{w}\partial_\xi \hat{v} = 0, \quad \forall v \in S_h.$$

By the Bramble-Hilbert Lemma,

$$|\int_{\hat{K}} \partial_\xi \hat{w}\partial_\xi \hat{v}| \leq C\|D^3\hat{u}\|_{L^2(\hat{K})}\|\partial_\xi \hat{v}\|_{L^2(\hat{K})} \leq C(h_K^{1+\alpha}|u|_{2,K} + h_K^2|u|_{3,K})|v|_{1,K}. \tag{3.18}$$

We have used (2.2) and (2.3) in the last step. Similarly,

$$|\int_{\hat{K}} \partial_\eta \hat{w}\partial_\eta \hat{v}| \leq C(h_K^{1+\alpha}|u|_{2,K} + h_K^2|u|_{3,K})|v|_{1,K}. \tag{3.19}$$

Next we discuss the cross terms. For any $v \in S_h$, we can express

$$\partial_\xi \hat{v} = \partial_\xi \hat{v}(0,0) + \eta\partial^2_{\xi\eta}\hat{v}, \quad \partial_\eta \hat{v} = \partial_\eta \hat{v}(0,0) + \xi\partial^2_{\xi\eta}\hat{v}.$$

Note that $\partial^2_{\xi\eta}\hat{v}$ is a constant. We write

$$\begin{aligned} &\int_{\hat{K}} (\partial_\xi \hat{w}\partial_\eta \hat{v} \pm \partial_\eta \hat{w}\partial_\xi \hat{v}) \\ = \;& \partial_\eta \hat{v}(0,0)\int_{\hat{K}} \partial_\xi \hat{w} \pm \partial_\xi \hat{v}(0,0)\int_{\hat{K}} \partial_\eta \hat{w} + \partial^2_{\xi\eta}\hat{v}(\int_{\hat{K}} \xi\partial_\xi \hat{w} \pm \int_{\hat{K}} \eta\partial_\eta \hat{w}). \end{aligned}$$

Since for $\hat{u} = \xi^2$, or $\hat{u} = \eta^2$,

$$\int_{\hat{K}} \partial_\xi \hat{w} = 0, \quad \int_{\hat{K}} \partial_\eta \hat{w} = 0.$$

Therefore, by the Bramble-Hilbert Lemma,

$$\begin{aligned}
&|\partial_\eta \hat{v}(0,0) \int_{\hat{K}} \partial_\xi \hat{w} \pm \partial_\xi \hat{v}(0,0) \int_{\hat{K}} \partial_\eta \hat{w}| \\
\leq\ & C\|D^3\hat{u}\|_{L^2(\hat{K})} \|\hat{\nabla}\hat{v}\|_{L^2(\hat{K})} \leq C(h_K^{1+\alpha}|u|_{2,K} + h_K^2|u|_{3,K})|v|_{1,K}. \qquad (3.20)
\end{aligned}$$

Next we consider,

$$\int_{\hat{K}} \xi \partial_\xi \hat{w} = \frac{1}{2}\int_{\hat{K}} (\xi^2-1)'\partial_\xi \hat{w} = -\frac{1}{2}\int_{\hat{K}} (\xi^2-1)\partial^2_{\xi^2}\hat{u},$$

$$\begin{aligned}
\partial^2_{\xi\eta}\hat{v} \int_{\hat{K}} \xi \partial_\xi \hat{w} &= -\frac{1}{2}\int_{\hat{K}} (\xi^2-1)\partial^2_{\xi^2}\hat{u}\partial^2_{\xi\eta}\hat{v} \\
&= \frac{1}{2}\int_{-1}^{1} (\xi^2-1)(\partial^2_{\xi^2}\hat{u}\partial_\xi \hat{v})(\xi,-1)d\xi \\
&\quad -\frac{1}{2}\int_{-1}^{1} (\xi^2-1)(\partial^2_{\xi^2}\hat{u}\partial_\xi \hat{v})(\xi,1)d\xi + \frac{1}{2}\int_{\hat{K}} (\xi^2-1)\partial^3_{\xi^2\eta}\hat{u}\partial_\xi \hat{v}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\partial^2_{\xi\eta}\hat{v} \int_{\hat{K}} \eta \partial_\eta \hat{w} &= \frac{1}{2}\int_{-1}^{1} (\eta^2-1)(\partial^2_{\eta^2}\hat{u}\partial_\eta \hat{v})(-1,\eta)d\eta \\
&\quad -\frac{1}{2}\int_{-1}^{1} (\eta^2-1)(\partial^2_{\eta^2}\hat{u}\partial_\eta \hat{v})(1,\eta)d\eta + \frac{1}{2}\int_{\hat{K}} (\eta^2-1)\partial^3_{\xi\eta^2}\hat{u}\partial_\eta \hat{v}.
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
&\partial^2_{\xi\eta}\hat{v}(\int_{\hat{K}} \xi \partial_\xi \hat{w} \pm \int_{\hat{K}} \eta \partial_\eta \hat{w}) \\
=\ & \frac{1}{2}\int'_{\partial\hat{K}} (t^2-1)\partial^2_t \hat{u}\partial_t \hat{v}dt + \frac{1}{2}\int_{\hat{K}} [(\xi^2-1)\partial^3_{\xi^2\eta}\hat{u}\partial_\xi \hat{v} \pm (\eta^2-1)\partial^3_{\xi\eta^2}\hat{u}\partial_\eta \hat{v}], \qquad (3.21)
\end{aligned}$$

where $\int'$ indicates a sign influence whenever it applies. For the second term on the right hand side of (3.21), we have, from (2.3),

$$\begin{aligned}
&\frac{1}{2}\int_{\hat{K}} [(\xi^2-1)\partial^3_{\xi^2\eta}\hat{u}\partial_\xi \hat{v} \pm (\eta^2-1)\partial^3_{\xi\eta^2}\hat{u}\partial_\eta \hat{v}] \\
\leq\ & C(h_K^{1+\alpha}|u|_{2,K} + h_K^2|u|_{3,K})|v|_{1,K}. \qquad (3.22)
\end{aligned}$$

In light of (3.18)–(3.22), we can express

$$\begin{aligned}
&\int_{\hat{K}} (\hat{\nabla}\hat{w})^T B^K \hat{\nabla}\hat{v} d\xi d\eta \\
=\ & \frac{b_{12}^K}{2}\sum_{j=1}^{4} |l_j|^2 \int_{l_j} (t(s)^2-1)\partial^2_s u \partial_s v ds + (O(h_K^{1+\alpha})|u|_{2,K} + O(h_K^2)|u|_{3,K})|v|_{1,K} \qquad (3.23)
\end{aligned}$$

where l_j are four sides of K. By the neighboring condition (2.13), any two adjacent elements K_1, K_2 that share a common edge satisfy (see Lemma 3.1)

$$\|B^{K_1} - B^{K_2}\| = O(h^\alpha),$$

Therefore, we have, by the trace theory,

$$\begin{aligned}
& |\frac{b_{12}^{K_1} - b_{12}^{K_2}}{2}|l|^2 \int_l (t(s)^2 - 1)\partial_s^2 u \partial_s v ds| \\
\leq\ & Ch^\alpha |l|^2 (h^{-1} \int_K |D^2 u D v| + \int_K |D^3 u D v + D^2 u D^2 v|) \\
\leq\ & C(h^{1+\alpha}|u|_{2,K} + h^2 |u|_{3,K})|v|_{1,K}. \qquad (3.24)
\end{aligned}$$

In the last step, we have used the inverse inequality. Adding up (3.23) with the edge integral estimated by (3.24), we obtain, under the Dirichlet boundary condition,

$$|\sum_{K\in\mathcal{T}_h} (\nabla w, A\nabla v)_K^*| \leq C(h^{1+\alpha}|u|_{2,\Omega} + h^2|u|_{3,\Omega})|v|_{1,\Omega}. \qquad (3.25)$$

b) we now consider $\int_{\hat{K}} \hat{v}\partial_\xi \hat{w}$ where we can express

$$\hat{v} = \hat{v}(0,0) + \partial_\xi \hat{v}(0,0)\xi + \partial_\eta \hat{v}(0,0)\eta + \partial_{\xi\eta}^2 \hat{v}\xi\eta.$$

Since for any $\hat{u} \in P_2(\hat{K})$, we have

$$\int_{\hat{K}} \partial_\xi \hat{w}(\hat{v}(0,0) + \partial_\eta \hat{v}(0,0)\eta) = 0,$$

by the same argument as in b), we have

$$|\int_{\hat{K}} \partial_\xi \hat{w}(\hat{v}(0,0) + \partial_\eta \hat{v}(0,0)\eta)| \leq C\|D^3\hat{u}\|_{0,\hat{K}}\|\hat{v}\|_{1,\hat{K}} \leq C(h_K^\alpha |u|_{2,K} + h_K|u|_{3,K})\|v\|_{1,K}. \qquad (3.26)$$

Next, by identities

$$\int_{\hat{K}} \partial_\xi \hat{w}\xi = -\frac{1}{2}\int_{\hat{K}} \partial_{\xi^2}^2 \hat{u}(\xi^2 - 1),$$

$$\int_{\hat{K}} \partial_\xi \hat{w}\xi\eta = -\frac{1}{4}\int_{\hat{K}} \partial_\xi \hat{w}(\xi^2 - 1)'(\eta^2 - 1)' = \frac{1}{4}\int_{\hat{K}} \partial_{\xi^2\eta}^3 \hat{u}(\xi^2 - 1)(\eta^2 - 1),$$

we have

$$\begin{aligned}
& |\int_{\hat{K}} \partial_\xi \hat{w}(\partial_\xi \hat{v}(0,0)\xi + \partial_{\xi\eta}\hat{v}\xi\eta)| \\
\leq\ & \|\partial_{\xi^2}^2 \hat{w}\|_{0,\hat{K}} \|\partial_\xi \hat{v}\|_{0,\hat{K}} + \|\partial_{\xi^2\eta}^3 \hat{w}\|_{0,\hat{K}} \|\partial_{\xi\eta}^2 \hat{v}\|_{0,\hat{K}} \\
\leq\ & C(h_K^\alpha |u|_{2,K} + h_K|u|_{3,K})|v|_{1,K}. \qquad (3.27)
\end{aligned}$$

Again, we have used (2.2), (2.3), and the inverse inequality in the last step.

Combining (3.26) and (3.27), we obtain

$$|\int_{\hat{K}} \hat{v}\partial_\xi \hat{w}| \le C(h_K^\alpha |u|_{2,K} + h_K |u|_{3,K})\|v\|_{1,K}. \tag{3.28}$$

Similarly, we have

$$|\int_{\hat{K}} \hat{v}\partial_\eta \hat{w}| \le C(h_K^\alpha |u|_{2,K} + h_K |u|_{3,K})\|v\|_{1,K}. \tag{3.29}$$

Note that $X_0^K = O(h_K)$, therefore,

$$|(\boldsymbol{b}_0 \cdot \nabla w, v)_K^*| = |\boldsymbol{b}_0 \cdot X_0 \int_{\hat{K}} \hat{v}\nabla \hat{w}| \le Ch_K(h_K^\alpha |u|_{2,K} + h_K |u|_{3,K})\|v\|_{1,K}. \tag{3.30}$$

Adding up all $K \in \mathcal{T}_h$ and using the Cauchy inequality, we obtain

$$|\sum_{K\in\mathcal{T}_h} (\boldsymbol{b}_0 \cdot \nabla w, v)^*| \le C(h^{1+\alpha}|u|_{2,\Omega} + h^2|u|_{3,\Omega})\|v\|_{1,\Omega}. \tag{3.31}$$

Combining (3.17), (3.25), and (3.31), we establish the assertion for $b_h(w, v)$. □

Theorem 3.3. Assume the same hypotheses as in Theorem 3.2. Let $u_h \in S_h$ be the finite element approximation of u that satisfies (3.1), and let $a(\cdot,\cdot)$ satisfy the discrete inf-sup condition on S_h. Then

$$|a(u - u_I, v)| \le C(h^{1+\alpha}|u|_{2,\Omega} + h^2|u|_{3,\Omega})\|v\|_{1,\Omega}, \tag{3.32}$$

$$\|u_h - u_I\|_{1,\Omega} \le C(h^{1+\alpha}|u|_{2,\Omega} + h^2|u|_{3,\Omega}). \tag{3.33}$$

Proof: Let $w = u - u_I$, and by Theorem 3.1,

$$\begin{aligned}
|a(w,v)_K - a_h(w,v)_K| &= |(\nabla w, A\nabla v)_K - (\nabla w, A\nabla v)_K^*| \\
&\le Ch_K^\alpha \|\nabla w\|_{0,K}\|\nabla v\|_{0,K}.
\end{aligned}$$

Adding all $K \in \mathcal{T}_h$ and using (2.14) with the Cauchy-Schwarz inequality, we have

$$|a(w,v) - a_h(w,v)| \le Ch^{1+\alpha}|u|_{2,\Omega}|v|_{1,\Omega}.$$

Recall Theorem 3.2, and we obtain

$$|a(w,v)| \le |a(w,v) - a_h(w,v)| + |a_h(w,v)| \le C(h^{1+\alpha}|u|_{2,\Omega} + h^2|u|_{3,\Omega})\|v\|_{1,\Omega},$$

which establishes (3.32). We then complete the proof by the inf-sup condition in

$$c\|u_h - u_I\|_{1,\Omega} \le \sup_{v\in S_h} \frac{a(u_h - u_I, v)}{\|v\|_{1,\Omega}} = \sup_{v\in S_h} \frac{a(u - u_I, v)}{\|v\|_{1,\Omega}} \le C(h^{1+\alpha}|u|_{2,\Omega_{1,h}} + h^2|u|_{3,\Omega}). \quad \square$$

4. Gradient Recovery.

In this section, we introduce and analyze a polynomial preserving recovery method (PPR). We define a gradient recovery operator $G_h : S_h \to S_h \times S_h$, on bilinear finite element space under a quadrilateral partition $\mathcal{T}_h$ in a following way: Given a finite element solution u_h, we first define $G_h u_h$ at all nodes (vertices), and then obtain $G_h u_h$ on the whole domain by interpolation using the original nodal shape functions of S_h.

Given an interior node (vertex) $\boldsymbol{z}_i$, we select an element patch ω_i, where

$$\bar{\omega}_i = \bigcup_{K \in \mathcal{T}_h, \bar{K} \cap \boldsymbol{z}_i = \boldsymbol{z}_i} \bar{K}.$$

We then denote all nodes on $\bar{\omega}_i$ (including O) as $\boldsymbol{z}_{ij}$, $j = 1, 2, \ldots, n (\geq 6)$, and fit a quadratic polynomial, in the least-squares sense, to the finite element solution u_h at those nodes. Using the local coordinates (x, y) with $\boldsymbol{z}_i$ as the origin, the fitting polynomial is

$$p_2(x, y; \boldsymbol{z}_i) = \boldsymbol{P}^T \boldsymbol{a} = \hat{\boldsymbol{P}}^T \hat{\boldsymbol{a}},$$

with

$$\boldsymbol{P}^T = (1, x, y, x^2, xy, y^2), \quad \hat{\boldsymbol{P}}^T = (1, \xi, \eta, \xi^2, \xi\eta, \eta^2);$$
$$\boldsymbol{a}^T = (a_1, a_2, a_3, a_4, a_5, a_6), \quad \hat{\boldsymbol{a}}^T = (a_1, ha_2, ha_3, h^2 a_4, h^2 a_5, h^2 a_6),$$

where the scaling parameter $h = h_i$ is the length of the longest element edge in the patch ω_i. The coefficient vector $\hat{\boldsymbol{a}}$ is determined by the linear system

$$Q^T Q \hat{\boldsymbol{a}} = Q^T \boldsymbol{b}_h, \tag{4.1}$$

where $\boldsymbol{b}_h^T = (u_h(\boldsymbol{z}_{i1}), u_h(\boldsymbol{z}_{i2}), \cdots, u_h(\boldsymbol{z}_{in}))$ and

$$Q = \begin{pmatrix} 1 & \xi_1 & \eta_1 & \xi_1^2 & \xi_1\eta_1 & \eta_1^2 \\ 1 & \xi_2 & \eta_2 & \xi_2^2 & \xi_2\eta_2 & \eta_2^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \xi_n & \eta_n & \xi_n^2 & \xi_n\eta_n & \eta_n^2 \end{pmatrix}.$$

The condition for (4.1) to have a unique solution is: Q has a full rank, which is almost always satisfied in practical situation when $n \geq 6$ and grid points are reasonably distributed. In fact, Q has a full rank if and only if there exits six nodes among $\boldsymbol{z}_{ij}$s that are not lying on a conic curve.

Now we define

$$G_h u_h(\boldsymbol{z}_i) = \nabla p_2(0, 0; \boldsymbol{z}_i). \tag{4.2}$$

When Neumann boundary condition is post, there is no need to do gradient recovery on the boundary. However, if the Dirichlet boundary condition is post, the recovered gradient on a boundary node $\boldsymbol{z}$ can be determined from an element patch ω_i such that $\bar{\omega}_i \cap \boldsymbol{z} = \boldsymbol{z}$ in the following way: Let the relative coordinates of $\boldsymbol{z}$ with respect to $\boldsymbol{z}_i$ is, say (h, h), then

$G_h u_h(\boldsymbol{z}) = \nabla p_2(h,h;\boldsymbol{z}_i)$. If $\boldsymbol{z}$ is covered by more than one element patches, then some averaging may be applied.

Remark. In an earlier work [24], Wiberg-Li least-squares fitted solution values to improve and to estimate the L_2-norm errors of the finite element approximation.

Now, we demonstrate PPR on an element patch that contains four uniform square elements (Figure 2). Fitting

$$\hat{p}_2(\xi,\eta) = (1,\xi,\eta,\xi^2,\xi\eta,\eta^2)(\hat{a}_1,\cdots,\hat{a}_6)^T$$

with respect to the nine nodal values on the patch. Now

$$\vec{e} = (1,1,1,1,1,1,1,1,1)^T, \quad \vec{\xi} = (0,1,1,0,-1,-1,-1,0,1)^T,$$

$$\vec{\eta} = (0,0,1,1,1,0,-1,-1,-1)^T, \quad Q = (\vec{e},\vec{\xi},\vec{\eta},\vec{\xi^2},\vec{\xi\eta},\vec{\eta^2}),$$

$$\begin{aligned}(Q^TQ)^{-1}Q^T &= \operatorname{diag}(\frac{1}{9},\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{4},\frac{1}{6})\cdot \\ &\quad \begin{pmatrix} 5 & 2 & -1 & 2 & -1 & 2 & -1 & 2 & -1 \\ 0 & 1 & 1 & 0 & -1 & -1 & -1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & -1 & -1 & -1 \\ -2 & 1 & 1 & -2 & 1 & 1 & 1 & -2 & 1 \\ 0 & 0 & 1 & 0 & -1 & 0 & 1 & 0 & -1 \\ -2 & -2 & 1 & 1 & 1 & -2 & 1 & 1 & 1 \end{pmatrix}.\end{aligned}$$

$$G_h u(O) = \frac{1}{h}\hat{\nabla}\hat{p}_2(0,0) = \frac{1}{6h}\begin{pmatrix} u_1 - u_5 + u_2 - u_4 + u_8 - u_6 \\ u_2 - u_8 + u_3 - u_7 + u_4 - u_6 \end{pmatrix} = \frac{1}{h}\sum_j \vec{c}_j u_j. \tag{4.3}$$

Note that the desired weights $\vec{c}_j$ are the second row of $(Q^TQ)^{-1}Q^T$ for the x-derivative, and the third row of $(Q^TQ)^{-1}Q^T$ for the y-derivative. Moreover, $\sum_i \vec{c}_j = \vec{0}$ and $G_h u(O)$ provides a second-order finite difference scheme at O.

Given $v \in S_h$, it is straightforward to verify that

$$\begin{aligned}
\frac{\partial v}{\partial x}(\frac{h}{2},\frac{2h}{3}) &= \frac{1}{3h}(v_1 - v_0) + \frac{2}{3h}(v_2 - v_3), \\
\frac{\partial v}{\partial x}(-\frac{h}{2},\frac{2h}{3}) &= \frac{1}{3h}(v_0 - v_5) + \frac{2}{3h}(v_3 - v_4), \\
\frac{\partial v}{\partial x}(-\frac{h}{2},-\frac{2h}{3}) &= \frac{1}{3h}(v_0 - v_5) + \frac{2}{3h}(v_7 - v_6), \\
\frac{\partial v}{\partial x}(\frac{h}{2},-\frac{2h}{3}) &= \frac{1}{3h}(v_1 - v_0) + \frac{2}{3h}(v_8 - v_7).
\end{aligned}$$

Therefore,

$$G_h^x v(O) = \frac{1}{4}[\frac{\partial v}{\partial x}(\frac{h}{2},\frac{2h}{3}) + \frac{\partial v}{\partial x}(-\frac{h}{2},\frac{2h}{3}) + \frac{\partial v}{\partial x}(-\frac{h}{2},-\frac{2h}{3}) + \frac{\partial v}{\partial x}(\frac{h}{2},-\frac{2h}{3})].$$

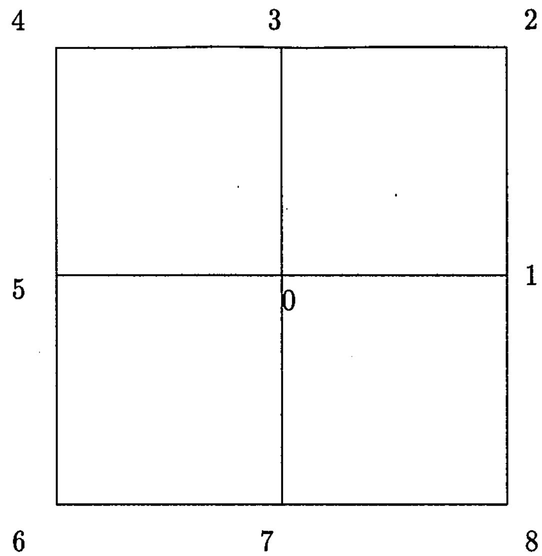

Figure 2

The recovered y-derivative can be obtained similarly. Hence, in this special case,

$$|G_h v(O)| \leq |v|_{1,\infty,\omega_O}.$$

By linear mapping, this is also valid for four uniform parallelograms in that

$$|G_h v(O)| \leq C|v|_{1,\infty,\omega_O}, \quad \forall v \in S_h. \tag{4.4}$$

with C independent of h and v.

Theorem 4.1 Let $\mathcal{T}_h$ satisfy *Condition* (α). Then the recovery operator G_h is a bounded linear operator on bilinear element space such that

$$\|G_h v\|_{0,p,\Omega} \leq C|v|_{1,p,\Omega}, \quad \forall v \in S_h, \quad 1 \leq p \leq \infty,$$

where C is a constant independent of v and h.

Proof: We observe that the diagonal condition together with the neighboring condition imply that for any given node O, there are four elements attached to it when h is sufficiently small. In addition, these four elements deviate from four parallelograms that attached to the same node in the following sense,

$$Q = Q_0 + h^\alpha Q_1,$$

where Q and Q_0 are least-square fitting matrices associated with those four quadrilateral elements and four parallelograms, respectively. We want to express $(Q^TQ)^{-1}Q^T$ in terms of $(Q_0^TQ_0)^{-1}Q_0^T$. Towards this end, we have

$$Q^TQ = Q_0^TQ_0(I + h^\alpha E_1),$$

where

$$E_1 = (Q_0^T Q_0)^{-1}(Q_1^T Q_0 + Q_0^T Q_1 + h^\alpha Q_1^T Q_1).$$

Therefore,

$$(Q^T Q)^{-1} Q^T = (I + h^\alpha E_1)^{-1}(Q_0^T Q_0)^{-1}(Q_0^T + h^\alpha Q_1^T) = (Q_0^T Q_0)^{-1} Q_0^T + h^\alpha E_2,$$

where

$$E_2 = (Q_0^T Q_0)^{-1} Q_1^T - \sum_{j=0}^{\infty} (h^\alpha E_1)^j E_1 (Q_0^T Q_0)^{-1} Q^T.$$

We see that

$$(Q^T Q)^{-1} Q^T = (Q_0^T Q_0)^{-1} Q_0^T + O(h^\alpha). \tag{4.5}$$

Therefore, the fact that $Q_0^T Q_0$ is invertible guarantees that $Q^T Q$ is invertible for sufficiently small h. Moreover, by (4.5), we have

$$G_h v(O) = \frac{1}{h} \sum_j (\vec{c}_j + O(h^\alpha)) v_j$$

where G_h is the recovery operator under the quadrilateral mesh that satisfies the *diagonal condition* and the *neighboring condition*, and $\vec{c}_j$s are weights for the related parallelogram mesh so that, by (4.4), $\frac{1}{h} \sum_j \vec{c}_j$ is a bounded operator on S_h such that

$$\left|\frac{1}{h} \sum_j \vec{c}_j v_j\right| \le C|v|_{1,\infty,\omega_O}.$$

Therefore, in the quadrilateral case, (4.4) is also valid provided *Condition* (α) is satisfied and h is sufficiently small. If (4.4) is valid for each node of K, then we have,

$$\|G_h v\|_{0,\infty,K} \le C|v|_{1,\infty,\omega_K}, \quad \forall v \in S_h, \tag{4.6}$$

where ω_K is defined as

$$\bar{\omega}_K = \bigcup_{K' \in \mathcal{T}_h, \bar{K}' \cap \bar{K} \neq \emptyset} \bar{K}'.$$

Note that (4.6) is true for all $K \in \mathcal{T}_h$ including boundary elements, since by our construction the boundary recovery is simply some averaging of nearby patches. Therefore,

$$\|G_h v\|_{0,\infty,\Omega} \le C|v|_{1,\infty,\Omega}, \quad \forall v \in S_h, \tag{4.7}$$

which has established the assertion for $p = \infty$. As for $p < \infty$, we notice that all norms are equivalent for finite dimensional spaces, and with a scaling argument,

$$\begin{aligned}
\sum_{K\in\mathcal{T}_h}\int_K |G_hv|^p &\leq C_1\sum_{K\in\mathcal{T}_h} h^2\|G_hv\|_{0,\infty,K}^p \\
&\leq C_2h^2\sum_{K\in\mathcal{T}_h}|v|_{1,\infty,K}^p \\
&\leq C_3h^2\sum_{K\in\mathcal{T}_h}h^{-2}\int_K|\nabla v|^p \leq C\sum_{K\in\mathcal{T}_h}|v|_{1,p,K}^p
\end{aligned}$$

Here, all constants C_j's and C are independent of p, v, and h. Conclusion then follows. □

Another important feature of the new recovery operator is the following polynomial preserving property.

Lemma 4.1. Let $K \in \mathcal{T}_h$ and u be a quadratic polynomial on ω_K. Assume that the least-squares matrix Q is of full rank. Then $G_hu = \nabla u$ on K.

Proof: The rank condition guarantees that the least-squares fitting has a unique solution. On each of four element patches, the recovery procedure results in a quadratic polynomial p_2 that least-squares fits u, a quadratic polynomial. Therefore, $p_2 = u$, and consequently, $G_hu = \nabla p_2 = \nabla u$, a linear function, at each of the four vertices of K. Therefore, $G_hu = \nabla u$ on K. □

Remark. Note that we do not make any mesh assumption in Lemma 4.1 except the rank condition. Basically, as long as the least-squares fitting procedure can be carried on, the polynomial preserving property is satisfied. As a comparison, the ZZ recovery operator does not have this polynomial preserving property under general meshes, see [29] for more details.

Theorem 4.2. Assume the same hypotheses as in Theorem 3.3. Then the recovered gradient is superconvergence in the sense

$$\|\nabla u - G_hu_h\|_{0,\Omega} \leq C(h^{1+\alpha}|u|_{2,\Omega} + h^2|u|_{3,\Omega}),$$

where C is a constant independent of u and h.

Proof: We decompose the error into

$$\nabla u - G_hu_h = \nabla u - G_hu + G_h(u_I - u_h). \tag{4.8}$$

Note that $G_hu = G_hu_I$ since $u_I = u$ at all vertices and the recovery operator G_h is completely determined by nodal values of u. By the polynomial preserving property and the Bramble-Hilbert lemma,

$$\|\nabla u - G_hu\|_{0,p,\Omega} \leq Ch^2|u|_{3,p,\Omega}, \quad 1 \leq p \leq \infty. \tag{4.9}$$

By Theorem 4.1, G_h is a bounded operator for all interior patches. Therefore,

$$\begin{aligned}\|G_h(u_I-u_h)\|_{0,\Omega}^2 &= \sum_{K\in\mathcal{T}_h}\|G_h(u_I-u_h)\|_{0,K}^2\\ &\le C^2\sum_{K\in\mathcal{T}_h}|u_I-u_h|_{1,K}^2 \le C^2(h^{1+\alpha}|u|_{2,\Omega}+h^2|u|_{3,\Omega})^2 \qquad (4.10)\end{aligned}$$

by Theorem 3.3. The conclusion then follows by applying (4.9) with $p=2$ and (4.10) to (4.8). □

5. A Posteriori Error Estimates.

Let $e_h=u-u_h$, the task here is to estimate the error $\|\nabla e_h\|_{0,\Omega}$ by a computable quantity η_h. According to Zienkiewicz-Zhu [32], η_h is the error estimator defined by the recovered gradient,

$$\eta_h=\|G_hu_u-\nabla u_h\|_{0,\Omega}.$$

We need the following assumption:

$$\|\nabla e_h\|_{0,p,\Omega}\ge Ch,\quad 1\le p\le\infty. \qquad (5.1)$$

Theorem 5.1. Assume the same hypotheses as in Theorem 4.2. Let (5.1) be satisfied with $p=2$. Then

$$\frac{\eta_h}{\|\nabla e_h\|_{0,\Omega}}=1+O(h^\rho),\quad \rho=\min(1,\alpha).$$

Proof: By the triangle inequality,

$$\eta_h-\|\nabla u-G_hu_h\|_{0,\Omega}\le\|\nabla e_h\|_{0,\Omega}\le\eta_h+\|\nabla u-G_hu_h\|_{0,\Omega}.$$

Dividing the above by $\|\nabla e_h\|_{0,\Omega}$, the conclusion follows from Theorem 4.2 and (5.1) with $p=2$. □

In order to estimate error on an interior element, we define

$$\eta_K=\|G_hu_h-\nabla u_h\|_{0,\infty,K}$$

and assume that

$$\|\nabla(u_I-u_h)\|_{L^\infty(\omega_K)}\le Ch^{1+\rho}. \qquad (5.2)$$

Theorem 5.2. Assume $u\in W^3_\infty(\Omega)$ and the same hypotheses as in Theorem 4.2. Let (5.1) and (5.2) be satisfied with $p=\infty$. Then

$$\frac{\eta_K}{\|\nabla e_h\|_{0,\infty,K}}=1+O(h^\rho).$$

Proof: By the triangle inequality,

$$\eta_K-\|G_hu_h-\nabla u\|_{0,\infty,K}\le\|\nabla e_h\|_{0,\infty,K}\le\eta_K+\|G_hu_h-\nabla u\|_{0,\infty,K};$$

and

$$\begin{aligned}\|G_h u_h - \nabla u\|_{0,\infty,K} &\leq \|G_h(u_I - u_h)\|_{0,\infty,K} + \|G_h u - \nabla u\|_{0,\infty,K} \\ &\leq C(\|\nabla(u_I - u_h)\|_{L^\infty(\omega_K)} + h^2|u|_{3,\infty,\omega_K}) \leq Ch^{1+\rho},\end{aligned}$$

by (4.9) and (5.2). The rest argument is the same as in Theorem 5.1. □

Remark 5.1. For Theorem 5.2, we made an assumption (5.2), which is the counterpart of (3.33) in the case of the L_∞-norm. To establish (5.2) usually involves delicate interior analysis, see Wahlbin [22] for more about interior analysis.

Remark 5.2. If we use $o(1)$ to substitute $O(h^\alpha)$, the right hand-side of Theorem 5.1 (or Theorem 5.2) would be $1 + o(1)$, the error estimate would be still asymptotically correct. Furthermore, we may use a more practical term "a sufficiently small constant $\gamma > 0$" instead of $o(1)$. Then we would lose the asymptotic exactness. Nevertheless, the effectivity index would still be in a reasonable range around 1. This is exactly what we have observed in practice.

Remark 5.3. We made an assumption that $\mathcal{T}_h$ satisfies *Condition* (α). Actually, all we need is that a large portion of K in $\mathcal{T}_h = \mathcal{T}_{1,h} \cup \mathcal{T}_{2,h}$ satisfy *Condition* (α) in the sense:

$$\bigcup_{K \in \mathcal{T}_{2,h}} \bar{K} = \bar{\Omega}_{2,h}, \quad |\Omega_{2,h}| = O(h^\sigma), \quad \text{with} \quad \sigma > 0,$$

and $\mathcal{T}_{1,h}$ satisfies *Condition* (α). This idea has been recently introduced by Bank-Xu [6, 7] for the strongly regular triangular grid, a counterpart $\alpha = 1$ in triangular mesh.

In addition, while all element K in $\mathcal{T}_{1,h}$ are quadrilaterals, an element K in $\mathcal{T}_{2,h}$ can be a triangle or a quadrilateral.

Appendix

Proof of Lemma 2.1. Let the longest edge length of K be h_K, then the longest edge length after one bisection refinement is $h_K/2$. We shall show that the distance between the two diagonal mid-points of any one of the four refined quadrilaterals is $d_K/4$, a quadratic reduction.

1) The coordinates O show that P_1P_3, P_4P_2 and O_1O_2 bisect each other at O.
2) $|O_1P_2| = |Z_3Z_4|/2 = |Z_3P_3|$ since O_1P_2 connects two edge centers in $\Delta Z_2Z_3Z_4$.
3) $|Z_3Q_1| = |Q_1O_1|$ since two triangles $\Delta Q_1O_1P_2$ and $\Delta Q_1Z_3P_3$ are congruent.
4) $|Q_1Q_2| = |OO_1|/2 = d_K/4$ since Q_1Q_2 connects two edge centers in ΔZ_3OO_1. □

Proof of Lemma 2.2. From

$$|\boldsymbol{v}_1||\boldsymbol{v}_3|\sin(\pi - \theta_1) = |\boldsymbol{v}_1 \times \boldsymbol{v}_3| = \frac{1}{2}|J_1^K| = \frac{1}{8}|P_4P_2 \times O_1O_2| = \frac{1}{8}|P_4P_2|d_K \sin\beta_K,$$

we have

$$\sin(\pi - \theta_1) = \frac{|P_4P_2 \times O_1O_2|}{8|\boldsymbol{v}_1||\boldsymbol{v}_3|} = \frac{|P_4P_2|d_K}{8|\boldsymbol{v}_1||\boldsymbol{v}_3|}\sin\beta_K.$$

Similarly,

$$\sin(\pi - \theta_2) = \frac{|P_1P_3|d_K}{8|\boldsymbol{v}_2||\boldsymbol{v}_4|} \sin \gamma_K.$$

Note that

$$\min(|\boldsymbol{v}_1|, |\boldsymbol{v}_3|) \le |P_4P_2| \le \max(|\boldsymbol{v}_1|, |\boldsymbol{v}_3|),$$
$$\min(|\boldsymbol{v}_2|, |\boldsymbol{v}_4|) \le |P_1P_3| \le \max(|\boldsymbol{v}_2|, |\boldsymbol{v}_4|);$$

and when σ_K is small,

$$\sin(\pi - \theta_1) \approx \pi - \theta_1, \quad \sin(\pi - \theta_2) \approx \pi - \theta_2.$$

The conclusion then follows. □

References

[1] G. Acosta and R.G. Durán, Error estimates for Q_1 isoparametric elements satisfying a weak angle condition, *SIAM J. Numer. Anal.* 38 (2001), 1073-1088.

[2] M. Ainsworth and J.T. Oden, *A Posteriori Error Estimation in Finite Element Analysis*, Wiley Interscience, New York, 2000.

[3] D.N. Arnold, D. Boffi, and R.S. Falk, Approximation by quadrilateral finite elements, *Math. Comp.* 71 (2002), 909-922.

[4] I. Babuška and T. Strouboulis, *The Finite Element Method and its Reliability*, Oxford University Press, London, 2001.

[5] R. Bank, Hierarchical bases and the finite element method, *Acta Numerica* (1996), 1-43.

[6] R.E. Bank and J. Xu, Asymptotically exact *a posteriori* error estimates, Part I: Grids with superconvergence, preprint.

[7] R.E. Bank and J. Xu, Asymptotically exact *a posteriori* error estimates, Part I: General unstructured grids, preprint.

[8] S.C. Brenner and L.R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer-Verlag, New York, 1994.

[9] Carsten Carstensen and Sören Bartels, Each averaging technique yields reliable *a posteriori* error control in FEM on unstructured grids. Part I: Low order conforming, nonconforming, and mixed FEM, *Math. Comp.* 71 (2002), 945-969.

[10] C. Chen and Y. Huang, *High Accuracy Theory of Finite Element Methods* (in Chinese), Hunan Science Press, P.R. China, 1995.

[11] Y. Chen, Superconvergent recovery of gradients of piecewise linear finite element approximations on non-uniform mesh partitions, *Numer. Meths. PDEs* 14 (1998), 169-192.

[12] P.G. Ciarlet, *The finite element method for elliptic problems*, North-Holland, Amsterdam, 1978.

[13] V. Girault and P.-A. Raviart, *Finite element methods for Navier-Stokes equations*, Springer-Verlag, New York, 1986.

[14] W. Hoffmann, A.H. Schatz, L.B. Wahlbin, and G. Wittum, Asymptotically exact *a posteriori* estimators for the pointwise gradient error on each element in irregular meshes. Part 1: A smooth problem and globally quasi-uniform meshes, *Math. Comp.* 70 (2001), 897-909.

[15] ~~W. Hoffmann,~~ A.H. Schatz, L.B. Wahlbin, ~~and G. Wittum,~~ Asymptotically exact *a posteriori* estimators for the pointwise gradient error on each element in irregular meshes. Part 2: The piecewise linear case, preprint.

[16] M. Křížek, P. Neittaanmäki, and R. Stenberg (Eds.), *Finite Element Methods: Superconvergence, Post-processing, and A Posteriori Estimates*, Lecture Notes in Pure and Applied Mathematics Series, Vol. 196, Marcel Dekker, Inc., New York, 1997.

[17] B. Li and Z. Zhang, Analysis of a class of superconvergence patch recovery techniques for linear and bilinear finite elements, *Numer. Meth. PDEs* 15 (1999), 151–167.

[18] Q. Lin and N. Yan, *Construction and Analysis of High Efficient Finite Elements* (in Chinese), Hebei University Press, P.R. China, 1996.

[19] P.B. Ming and Z.-C. Shi, Quadrilateral mesh revisited, preprint.

[20] R. Rannacher and S. Turek, Simple nonconforming quadrilateral Stokes element, *Numer. Meth. PDEs* 8 (1992), 97-111.

[21] R. Verfürth, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, Wiley-Teubner, Stuttgart, 1996.

[22] L.B. Wahlbin, Local behavior in finite element methods, in *Handbook of Numerical Analysis, Vol. II, Finite Element Methods* (Part 1), P.G. Ciarlet and J.L. Lions, ed., Elsevier Science Publishers B.V. (North-Holland), 1991, 353-522.

[23] L.B. Wahlbin, *Superconvergence in Galerkin Finite Element Methods*, Lecture Notes in Mathematics, Vol. 1605, Springer, Berlin, 1995.

[24] N.-E. Wiberg and X.D. Li, Superconvergence patch recovery of finite element solutions and *a posteriori* L_2 norm error estimate, *Commun. Num. Meth. Eng.*, 37 (1994), 313–320.

[25] J. Xu and Z. Zhang, Analysis of Recovery Type A Posteriori Error Estimators for Mildly Structured Grids, Research Report 2002 Series No. 7, Department of Mathematics, Wayne State University.

[26] N. Yan and A. Zhou, Gradient recovery type *a posteriori* error estimates for finite element approximations on irregular meshes, *Comput. Methods Appl. Mech. Engrg.* 190 (2001), 4289-4299.

[27] J. Zhang and F. Kikuchi, Interpolation error estimates of a modified 8-node serendipity finite element, *Numer. Math.* 85 (2000), 503-524.

[28] Z. Zhang, Analysis of some quadrilateral nonconforming elements for incompressible elasticity, *SIAM J. Numer. Anal.* 34-2 (1997), 640-663.

[29] Z. Zhang and A. Naga, A meshless gradient recovery method, Part 1: Superconvergence property, Research Report 2002 Series No. 2, Department of Mathematics, Wayne State University.

[30] Q.D. Zhu and Q. Lin, *Superconvergence Theory of the Finite Element Method* (in Chinese), Hunan Science Press, China, 1989.

[31] O.C. Zienkiewicz and J.Z. Zhu, The superconvergence patch recovery and *a posteriori* error estimates. Part 1: The recovery technique, *Int. J. Numer. Methods Engrg.* 33 (1992), 1331-1364.

[32] O.C. Zienkiewicz and J.Z. Zhu, The superconvergent patch recovery and *a posteriori* error estimates, Part 2: Error estimates and adaptivity, *Int. J. Numer. Methods Engrg.* 33 (1992), 1365-1382.

[33] M. Zlámal, Superconvergence and reduced integration in the finite element method, *Math. Comp.* 32 (1978), 663-685.

Department of Mathematics, Wayne State University, Detroit, MI 48202.
E-mail address: zzhangmath.wayne.edu